DOCUMENT RESUME

ED 222 519 TM 820 665

AUTHOR Noggle, Nelson L.
TITLE Supplementing Summative Findings with Formative Data.
PUB DATE 19 Mar 82
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (66th, New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Data Collection; Elementary Secondary Education; *Evaluation Methods; Field Studies; *Formative Evaluation; Migrant Education; Migrant Programs; *Summative Evaluation
IDENTIFIERS Aggregation (Data); California Assessment Program; *California Statewide Evaluation of Migrant Educ; *Evaluation Reports

ABSTRACT

This paper attempts to provide evaluators, administrators, and policy makers with the advantages of and methodology of merging formative and summative data to enhance summative evaluations. It draws on RMC Research Corporation's 1980-81 California Statewide Evaluation of Migrant Education. The concern that evaluations typically fail to obtain the "real story" was the underlying challenge for the approach used. The aggregate of 22 separate formative field studies was merged with the typical summative data base. The formative studies were compelling qualitative efforts in and of themselves; hence, they provided excellent sources for fleshing out the skeletal results from the summative effort. As a result, practitioners at all levels have accepted and are using the evaluation findings. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED222519

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SUPPLEMENTING SUMMATIVE FINDINGS WITH FORMATIVE DATA

Nelson L. Noggle

RMC Research Corporation
2570 West El Camino Real
Mountain View, CA 94040

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

N. L. Noggle

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented At
1982 AERA Annual Meeting
New York City

March 19, 1982

Tm 820.665

Preface

This paper attempts to provide evaluators, administrators, and policy makers with the advantages of and methodology of merging formative and summative data to enhance summative evaluations. It draws on RMC's 1980-81 California Statewide Evaluation of Migrant Education. The concern that evaluations typically fail to obtain the "real story" was the underlying challenge for the approach used. The aggregate of 22 separate formative field studies was merged with the typical summative data base. The formative studies were compelling qualitative efforts in and of themselves; hence, they provided excellent sources for fleshing out the skeletal results from the summative effort. As a result, practitioners at all levels have accepted and are using the evaluation findings.

Supplementing Summative Findings with Formative Data

Introduction

The educational evaluator, or researcher, customarily chooses among qualitative, quantitative, formative, and summative contexts within which to design her or his methodology. Such things as selecting instruments, sampling units of analysis, and choosing analysis strategies depend heavily upon which combination of these contexts controls the orientation of the study.

A major concern which stimulated this paper is that none of these decisions may have anything whatsoever to do with the planning, implementation, and outcomes of educational programs, materials, or strategies. These decisions may not affect education at all. No matter how rigorous the method, or how appropriate the context, these choices may matter only to evaluators themselves. The educational decision-makers tend to be only concerned with the "sound" of the evidence. If the results sound reasonable, the decision-maker chooses to use them to his or her best advantage. If the results sound unreasonable, the decision-maker pokes at method and context until she or he finds something missing, or amiss. The educational decison-makers can always find something wrong with the methods or context of an evaluation; but they will only do so when the results are surprising or especially contradictory to their beliefs or wishes.

If evaluators depend primarily upon quantiative and summative orientations to their study, they run the risk of missing the "real story" behind the results. They are left to their own experiences, or the experiences of consultants, to provide the interpretation that is so badly needed. Or, they leave the study at a descriptive level; and escape interpretive errors by leaving interpretation to the readers of the evaluation report, i.e., to the same decision-makers who react primarily on the sound of the evidence. This is unfortunate. When decision-makers are responsible for interpretation, they tend to interpret in the direction of their beliefs.

However, if evaluators orient their study only in a qualitative and formative context, they run the risk of being accused of not relying on strong statistically significant techniques. The decision-makers have been ritualized by their training to expect quantifiable results based on rigorous inferential statistical approaches involving representative samples. More likely, they may not fund qualitative or formative evaluations in the beginning because of costs. Decision-makers may not be able to cost justify a study that does not meet their understanding of appropriate evaluation methodology.

But the evaluators are faced with an evermore pressing problem. Decision-makers are often too busy to take the time to understand and use large or complex evaluation projects; they only want the "bottom-line." This mentality is in direct conflict with the needs of program participants, the people "in the trenches" trying to make the educational program work. They need information from an evaluation that keeps with day-to-day decisions about subcomponents of the program. If we plan around their needs, the decision-maker is left with very little. If we plan for the decision-maker, program people are left out.

To be good, an evalution must address the needs of those associated with the program, and it must be done in a fair and interpretable manner. This paper shares some insights about an evaluation project that has attempted to merge the various contexts for evaluation methods, both qualitative and quantitative, plus both formative and summative. The project was not as successful as it could have been; but, it provided the participating evaluators and decision-makers ample vision to have hope that such projects can and should be done.

The Evaluation Plan

The evaluation of migrant education in California is being conducted by the RMC Research Corporation, under contract with the California Department of Education. During the study three very important issues had to be addressed by the evaluation methodology: (1) how to

obtain the "real story" statewide, (2) how to evaluate a complex multi-level supplementary program providing educational and health services to an extremely mobile student population, and (3) how to provide the people in the trenches with feedback that helps them improve services to students, and at the same time provide administrators and policy makers with information that helps with all aspects of running the program. The State's Office of Program Evaluation and Research (OPER) knew that this was a problem and asked for an evaluation that combined the qualitative richness of local-level formative studies with the quantitative rigor of a statewide summative study.

The Evaluation Process

The evaluation of California's migrant education is in its second year. What is described here happened during the first year. This year's evaluation plan was changed somewhat based on the results of last year's experiences.

The district (13) and regional (9) field-sites were visited at three different times. The first visit was primarily a planning visit, and accomplished three things: (1) established rapport with the field-site personnel, (2) collected relevant planning documents and records, and (3) obtained input about the content and scope of both levels of the evaluation, the field-site and statewide. A totally unstructured interview technique was used during the first site visits.

Input from the first visits was integrated with direction from the State, and advice from the study's Advisory Panel. The result was a complete listing of the evaluation questions addressing all components of the program at all levels. The second and third site visitations were comprised primarily of semi-structured interviews centering around the evaluation questions, plus systematic reviews of the site's documents and records. Unstructured observation accounted for the remaining field activity.

Based on the evaluation questions generated early in the study, summative questionnaires were developed for the various personnel responsible for delivering migrant education services, plus migrant parents and students, and sent to 230 migrant education districts. Half of the districts were asked about 1980-81, and half were asked about 1979-80. Except for parents whose return rate was only 34%, the average return rate was 65%. Statistical checks for bias between years and waves of questionnaires were conducted and rejected at 0.05 level.

Since the State was against the additional burden of testing, certain achievement data were obtained through forms requesting grade advancement, graduation, and proficiency exam results from 117 districts. The return rate of these data was 75% (this was California's first year of proficiency testing, and districts were asked only to pilot-test their instruments). Grade-three achievement data from the California Assessment Program (CAP) was also accessed, since migrant students had been coded separately. Additional enumerative data were collected through access to Migrant Education monthly enrollment records, and Migrant Student Record Transfer System (MSRTS) records.

The evaluation team met following the site visits in a series of meetings to discuss, compare, and aggregate findings. Evaluation team members responsible for particular migrant education components led each meeting and were responsible for recordkeeping for that meeting. The focus of data aggregation surrounded the original evaluation questions to be answered, plus whatever new questions emerged during the site visits.

In addition, each member of the evaluation team was responsible for writing a final formative report which summarized the findings from each site they visited. Each team-member was responsible for four to six site reports, depending on their other responsibilities. Each of the written reports, as well as the earlier oral reports after each site visit (each site was visited three times), were provided directly to the person responsible for directing the Program at that site. Additional

people were provided these information only upon the direction of this person. The strictest of anonymity and confidentiality were followed throughout the course of the evaluation.

In the final analysis, 69 formative oral reports were given during the evaluation, and 23 written reports were provided. These figures include formative and summative reports to the State.

Merging the Formative with the Summative

The unique thing about this evaluation was that it actually used a semi-structured case-study approach in each field site, observing, interviewing, reviewing documents, and gathering enumerative data, while gathering questionnaire, achievement, and other enumerative data from samples large enough and representative enough for a Statewide summative report covering a specified time period. In other words, the evaluation accomplished what it set out to do. Hypotheses or conclusions which were derived from the summative data, were embellished, confirmed, or countered by the findings that emerged from the field studies. Conversely, hypotheses or tentative conclusions which seemed to emerge from the field work, could often be substantiated or rejected through post hoc analyses of the summative data. The majority of both sets of data, however, was formed through the evaluation questions asked at the beginning of the study.

On the good side. This paper will not go into the rich intercourse between evaluation team members which continually refined the field-site approach, and which ultimately helped each team member with their specific site reports. The aim of this paper is to demonstrate where the formative studies were and were not helpful with the Statewide summative report, and why.

The first, and perhaps greatest impact of the formative field effort, was in the formation of the summative evaluation plan. The summative plan was not finalized until after the evaluation team had made

their first round of site visits. It was at this time that the Advisory Panel was convened and offered the major thrust of the evaluation approach. Many of the probing networks of interview and questionnaire items grew out of the discourse following the first round of visits. Whereas the larger more general questions could be obtained from the State, the subtle intricate uses of detailed probes were derived from the feel of what was actually happening or needed by program participants in the field. If the real story was what the evaluators were after, it continually emerged with increased close contact with the field.

The second impact of the formative was the formulation of a series of hypotheses or counterquestions throughout the course of the data collection phase. As team members talked after each site visit, pieces of the puzzle began to fit together, and an understanding of how the program operates, began to grow.

The third impact of the formative was the growing rapport between the evaluators and program participants in the field. A mounting majority of those contacted, stated increased interest in our findings and faith that the findings would be believable. They knew that the results, since they were to get their own report on their own project, would be useful. However, as time went on, they pressured us to make sure the State and other program administrative levels would pay attention to the Statewide results. In a sense, the field began to support and take ownership in the Statewide summative evaluation as well as their own local evaluation.

The fourth impact came during the analysis of summative data. Three things occurred: (1) specific sets of questionnaire data were included or excluded because of the now known relevance to what was actually occurring in the field; (2) data layouts were often formated or centered around questions arising from the field work; and (3) the findings indicated by the summative data could be traced among the formative data for confirmation or clarification. In addition, the Advisory Panel

was reconvened to see initial (draft) layouts of the summative data. The Panel's questions or concerns prompted discussions which almost invariably required use of formative findings.

Finally, the fifth impact, and perhaps the most observable, was the use of formative data during the writing of the summative report. Several things happened during this stage. The report was initially written in separate components, with each component written by the evaluation team member responsible for that component. Based on the meetings that were held to discuss and aggregate the formative data, each team member drafted their components using appropriate portions of the aggregated formative data. The initial drafts were then circulated to the other team members, who proofread the sections in terms of their understanding from the data, both formative and summative. The basic result of this proofreading led to a sort of validation that the formative analysis conducted earlier had reached a nearly consensus level, and that the report made reasonable references to the formative data. When one looks at the report, one senses that the interpretations of data and the subsequent conclusions and recommendations had been grounded in program reality, and not in the "ivory-tower."

On the bad side. A major problem throughout the study was the fact that the formative studies were being conducted by five different evaluators--how could we cope with individual differences in style and interests, not to mention biases? To some degree, the staff meetings, referred to as "crosstalks," achieved certain levels of conformity, both in approach to data collection and in approach to analysis and reporting. The semi-structured interview guides and list of documents to be reviewed which arose out of the basic set of evaluation questions, added some additional commonality. Staff training, too, helped to secure more consistency across evaluators. And, during the analysis and reporting, the evaluators had a chance to delete, question, or add portions of the building story. Despite all of these steps, everyone of the evaluators felt uneasy and concerned about differences among themselves, such as the varying focus of extra probing in the field, the different levels of

knowledge in certain areas to be investigated, and diverse styles of working with field contacts. Unfortunately, the time and budget for the evaluation were too tight to allow teams of evaluators, or to allow the project director to monitor field work directly.

Another problem came from a basic, almost philosophical, difference among evaluators regarding how much faith to place in the formative data since it came from a small case-study sample of sites. This particular problem surfaced again and again, and was a major stumbling block at one point during report writing. Some of the evaluators could not place enough reliance on the rich data from such a small, perhaps nonrepresentative, sample. They tended to be more afraid to embellish, confirm, clarify, or reject summative results based on formative data. They were not afraid, however, to raise hypotheses based on formative data, and use summative data to verify or reject. The summative data base, however lacked flexibility since it was a fixed set of questions and information. Gaps in the summative data collection effort remained gaps forever; whereas, the formative data collection allowed the real story to emerge, even if the evaluation probes were faulty to begin with. This particular dilemma was basically resolved through the circulation of component drafts of the report, and through occasional resolutions by the project director.

The final problem rested with the level of ownership of the evaluation by Program administrators. As mentioned in the introduction, administrators are often too busy to involve themselves in the details and methodology of an evaluation effort. From the evaluators' point-of-view, had there been a higher level of ownership, the formative effort would have dove-tailed with the summative analysis much better than it did. The summative report was hard-hitting and rather assertive, and unfortunately it took certain administrators by surprise, which reflects badly on the formative. Why?

The evaluation team was not able, in the final analysis, to achieve the degree of rapport at the top level as it had with lower levels of

the Program. There was more ownership in the field than at the State. And, it could have been a major drawback in subsequent evaluation use. Such was not the case. The State-level personnel tended to somehow respect the report, and the intent of the evaluation team. The major findings of the report were focused upon, and there have been several policy or administrative decisions that addressed findings or recommendations from the report. In fact, the use made of the report has spurred the evaluators to realize that if they had improved rapport at the top, even more use may have been made of it. They have resolved to critically review how they as evaluators can do a better job of involving top administrators.

Summary

The summative report would have been little more than an enumerative report without the use of formative data. The formative data rose above the usual qualitative data since the formative data collection was designed to provide formative information to field participants. The fact that field participants could see utility for themselves, caused them to take more ownership in providing the "real story." As a result, they grew to understand and respect the evaluators sense of the mounting story and also grew to take ownership in the Statewide study. This led to a rich data base that could be merged with summative data and help create a summative report that was based on Program reality. There was usefulness in the report. There was use made of it.